
HERITAGEN: Next-Generation Sequencing and supercomputing for the unification of the genealogical and genetic heritage of Extremadura. Application to the study of hereditary diseases.

Investigadores:

- José Luis González Sánchez (Principal Investigator). COMPUTAEX Foundation.
- Silvia Romero Chala. San Pedro de Alcántara Hospital.
- Jonathan Gómez Raja. FUNDESALUD (Foundation for Research and Training of Health Professionals of Extremadura).
- José Antonio García Trujillo. San Pedro de Alcántara Hospital.
- Felipe Lemus Prieto. COMPUTAEX Foundation.
- Alfonso López Rourich. COMPUTAEX Foundation.

Idioma Inglés

Descripción:

It is estimated that approximately 1 in 200 births can be affected by the 6,000 known monogenic diseases. Establishing the pathogenicity of the mutations detected by means of NextGeneration Sequencing in the sequence of the genes involved will be of vital importance for the development of therapies in order to develop the Precision Medicine concept.

However, the lack of access to information about these variations means that their pathogenicity is unknown (VUS, V ariant of Uncertain Significance). The percentage of VUS can be reduced by accessing as much information as possible about the gene related to the disease under investigation, where the variations found in it are particularly important. The main problem is that the information is scattered, which hinders the access to it and supposes an economic expense and an increase of the difficulty in processing information.

One of these sources of information is the genealogical one, which is very useful in the study of the incidence of hereditary diseases within a family, in consultations of genetic counseling. However, the documentary heritage of the populations can also be found scattered, with limited or no access, so the genealogical information is limited to that provided by patients when they request a consultation.

The objective of the project is to study the benefits of the unification of heterogeneous information sources to the study of hereditary diseases (specifically the genealogical heritage and genetic information), which will serve to reduce the ratio of variables of uncertain significance detected in next-generation sequencing studies. To do this, it is proposed to focus the study on a group of people from a genetically relevant population, in order to sequence their genome (specifically the genes associated with the chosen disease) and to unify it with their genealogical information.

To manage and analyze the information, as well as to extract knowledge of it, the use of supercomputing resources in combination with big data techniques is proposed. Thus, by using HPC (High Performance Computing) techniques, the data and genetic sequences will be processed in the shortest possible time, while guaranteeing storage, security and high availability of the information used.

In addition, the use of standardized forms of information will be taken into account, as well as the ethical aspects derived from the performance of health-related treatments.

The high-level information generated will follow the open data philosophy, always complying with current legislation on security and information protection. Likewise, the data will be made available to users through services deployed by means of the cloud computing paradigm.

Fuentes de financiación:

- Project co-financed by the Junta of Extremadura and the European Regional Development Fund (ERDF) of Extremadura at 80 percent, within the Thematic Objective 01 "Reinforcement of research, technological development and innovation", through the call for grants for the realization of research projects, oriented towards the strategic areas of the regional economy contemplated in the V Regional R+D+i Plan (2014-2017), in public R+D+i centers of the Autonomous Community of Extremadura, under Decree 68/2016 of June 6.

URL del envío:<https://web.computaex.es/node/1603>